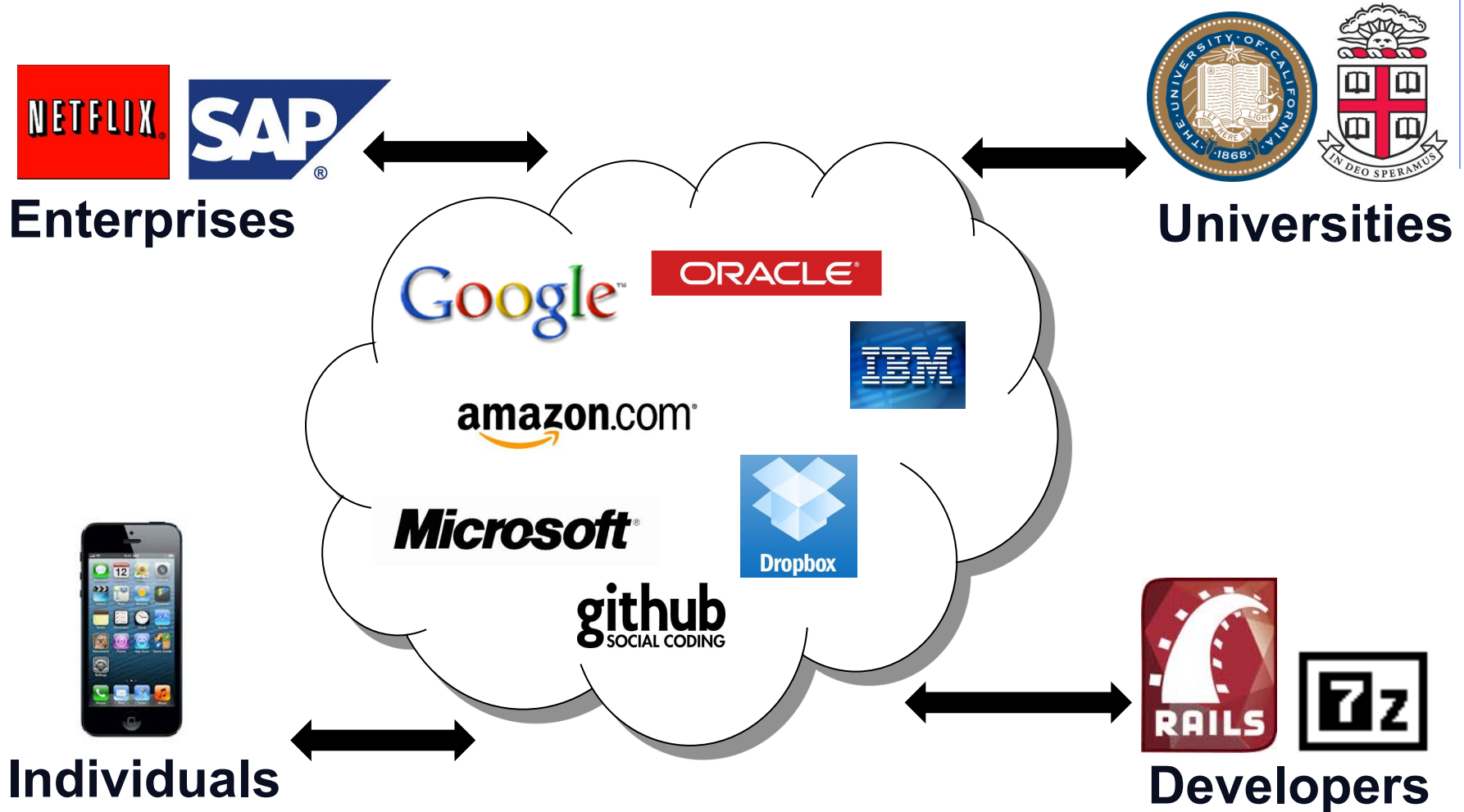


Secure Cloud Storage

Charalampos (Babis) Papamanthou
ECE and UMIACS
University of Maryland
cpap@umd.edu



Cloud computing today



Are there any threats?

- Cloud providers are untrusted
 - Can **lose** data
 - Can return **corrupted** results
 - Can **leak** information

...we will have no liability to you for any unauthorized access or use, corruption, deletion, destruction or loss of any of your content or applications...

Amazon web service customer agreement
<http://aws.amazon.com/agreement/>

WEB & COMMUNICATION SOFTWARE security
Hotmail Data Loss Reveals Cloud Trust Issues
Jan 3, 2011 11:56 AM

By Keir Thomas, PCWorld

News

Amazon struggles to restore lost data to European cloud customers

Developers vent frustration on Amazon support forum

By Jon Brodwin, Network World
August 09, 2011 11:17 AM ET

Gmail Corrupting Attachments

I recently received a report that attachments sent to Gmail from some servers

« [Security Recommendation](#) | [Main](#) | [Solaris Security](#)

Amazon S3 Silent Data Corruption

By user12606733 on Jan 28, 2009

While catching up on my reading, I came across an inter...

01 August 2012, 12:39

Dropbox confirms data leak

Cloud storage service provider [Dropbox](#) has [acknowledged](#) that a file

BPOS: a data leak in Microsoft's cloud

December 28th, 2010 - 09:10 am ET by J. G.

A configuration error in Microsoft's Business Productivity

Do people care?

- Customers are paying for the services
 - They want **reliable** storage
 - They want **correctness** guarantees
 - They want to keep their **privacy**

...58% of the public and 86% of business leaders are excited about the possibilities of cloud computing. But more than 90% of them are worried about security, availability, and privacy of their data as it rests in the cloud...

Microsoft survey in 2010

http://news.cnet.com/8301-1009_3-10437844-83.html

What do I do?

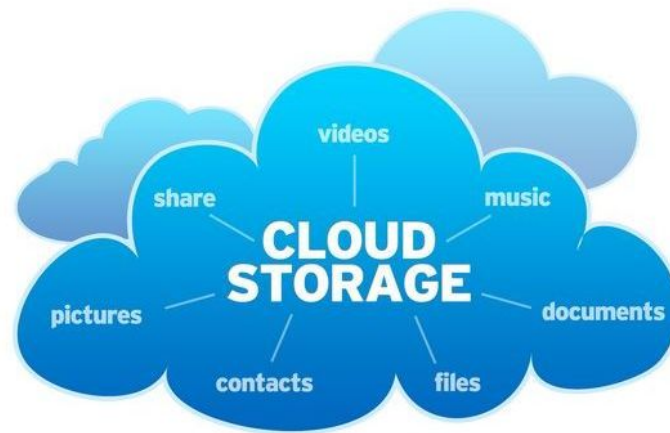
- Enable people to use the cloud safely
- Verifiability in the cloud
 - **Verify** that the cloud did the work correctly
- Privacy in the cloud
 - Use the cloud in a **privacy-preserving** manner

efficient both in theory and in practice

provably secure

no assumptions at the server

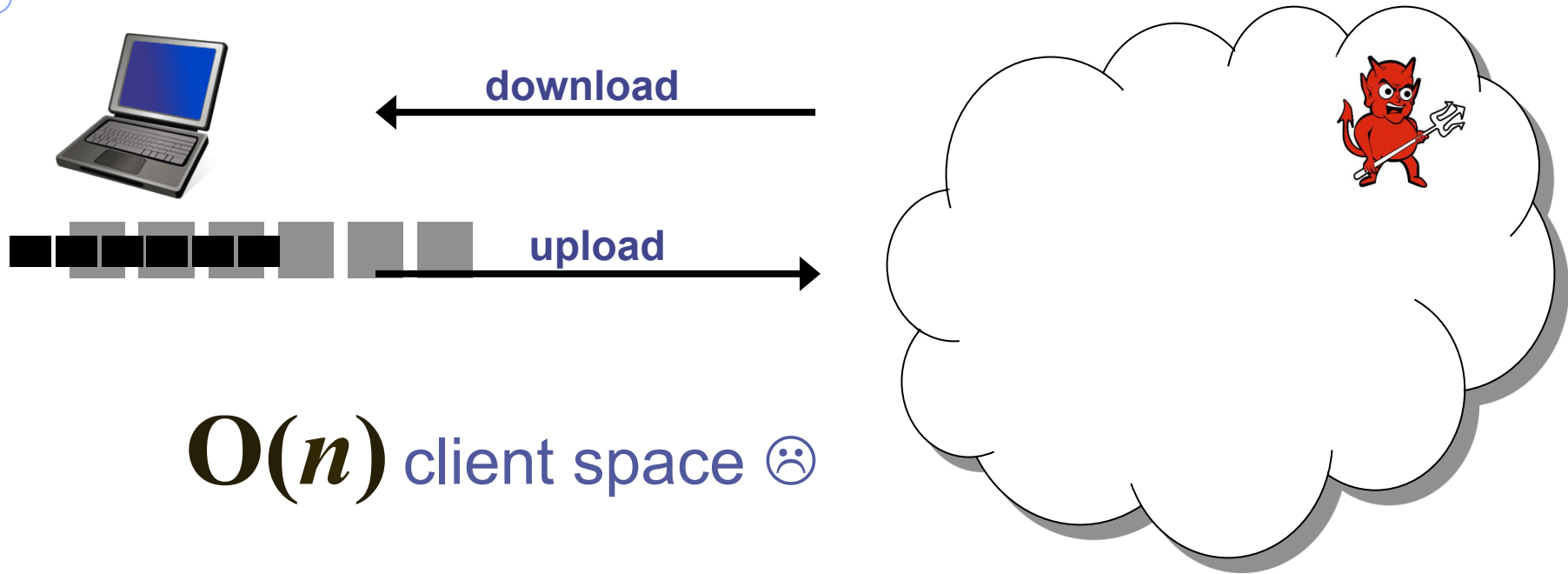
Secure Cloud Storage



Security Framework

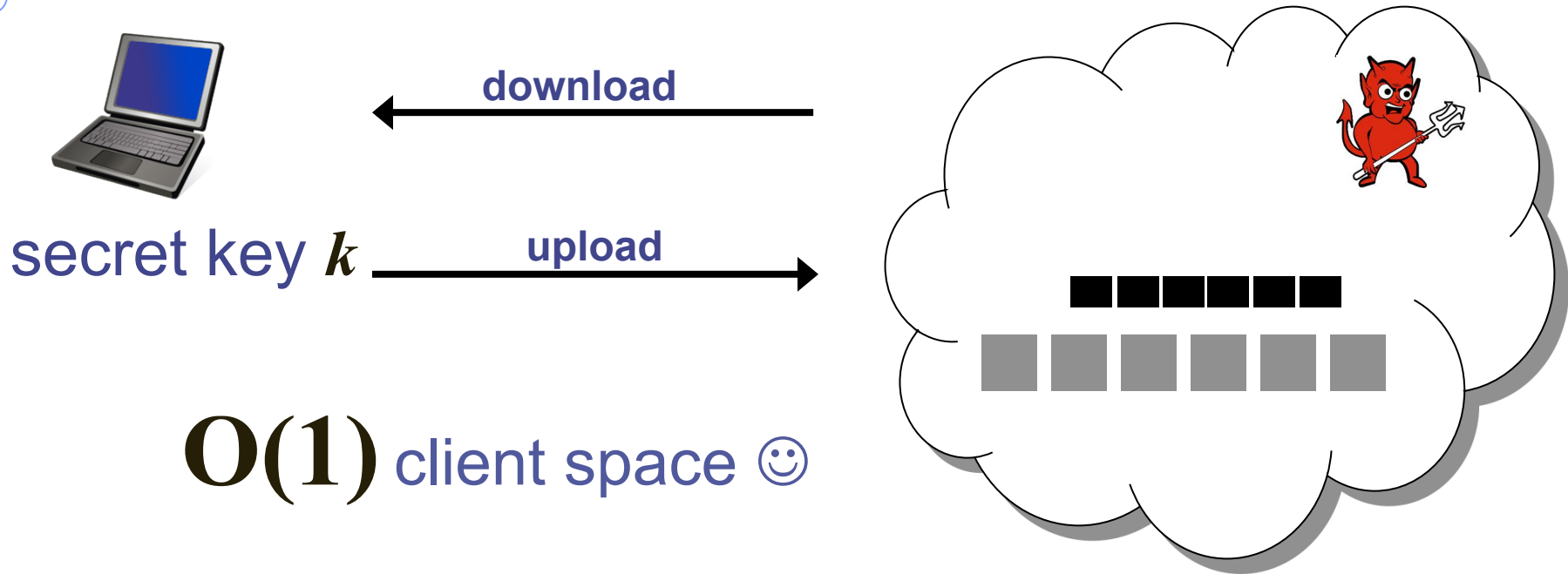
- We must make sure our files have not changed since they were uploaded
- We are going to ask the server that stores our files to compute a “proof” that he stores our files intact
- Central to the rest of the talk:
 - Cryptographic hash function, e.g., SHA256
 - Let's try it out

Storing your files in the cloud: Hash-based



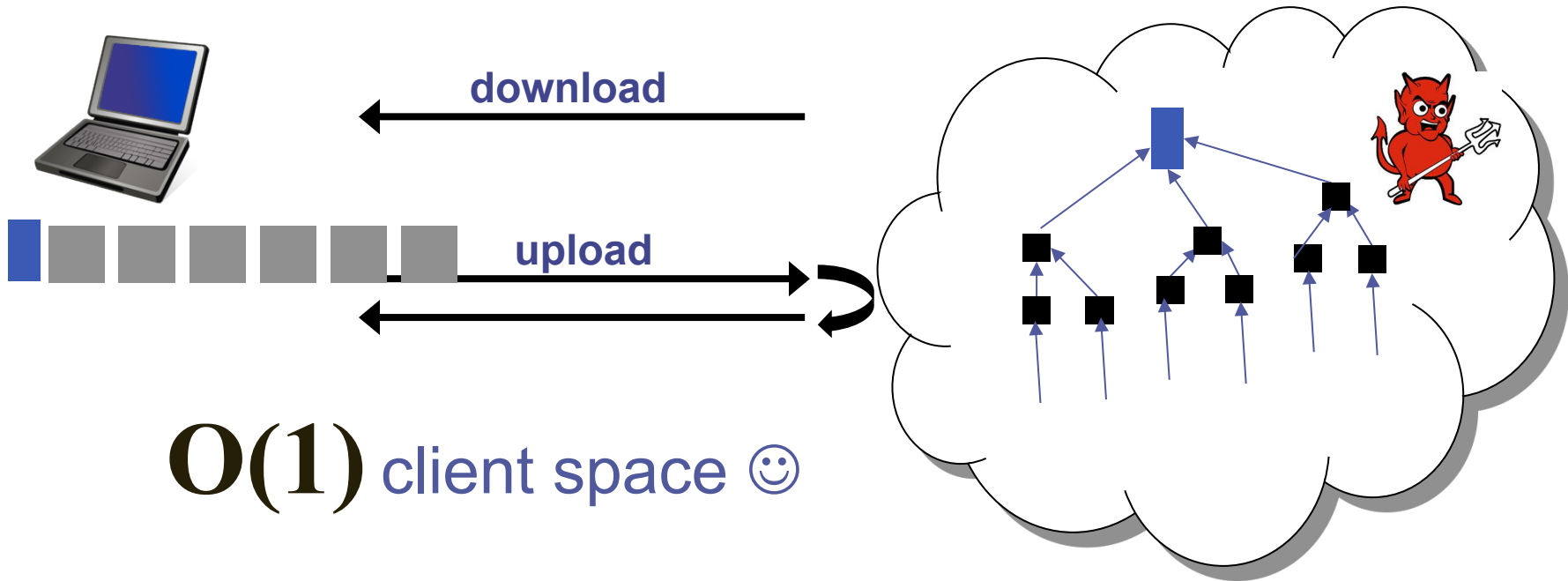
- How to verify that a file has **not** been corrupted?
 - Keep a hash (i.e., checksum) **locally** for each file
 - Download: **recompute** and **check**
 - Upload: **compute** and **store** new hash

Constant space? MAC-based



- How to verify that a file has **not** been corrupted?
- Compute a MAC for each file using a secret key
 - Store only the secret key!
 - Download: **recompute** and **check**
 - Upload: ?

What about replay attacks? Tree-based



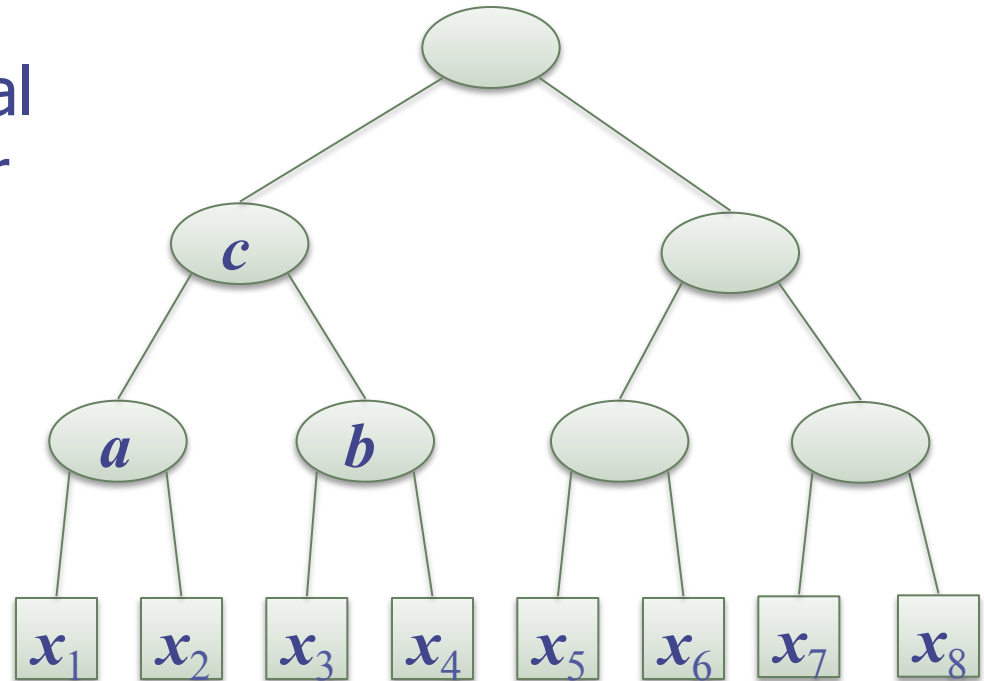
- Hashing over a tree and store only the roothash
 - Download: Fetch $O(\log n)$ hashes
 - Upload: An **interactive** protocol

Hash Tree: Details

- Balanced binary tree defining a hierarchical hashing scheme over a set of items

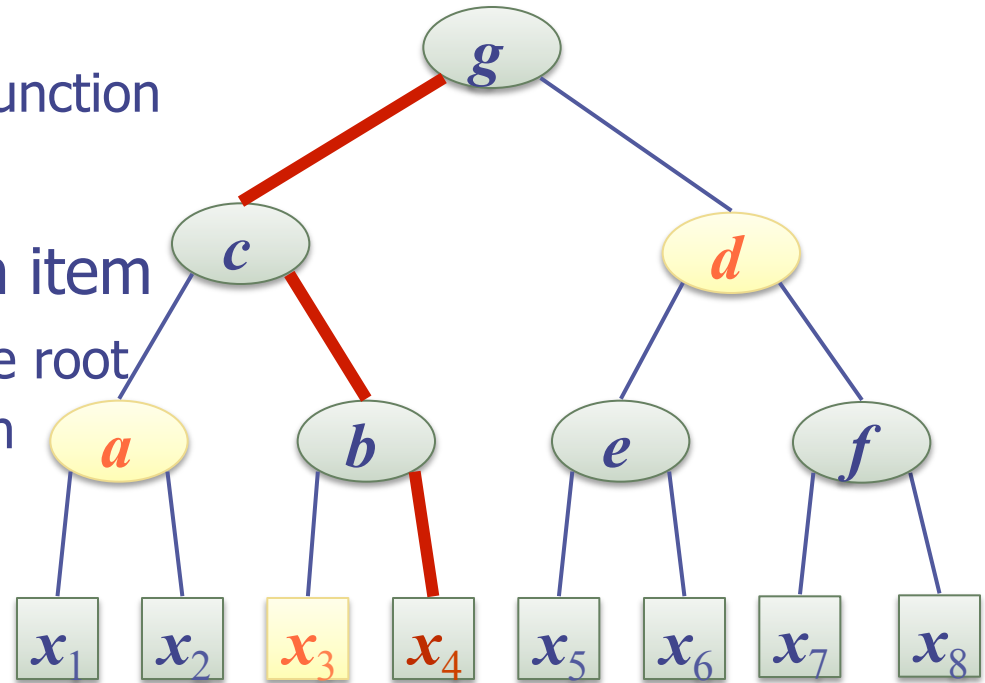
- $a = h(x_1, x_2)$
- $b = h(x_3, x_4)$
- $c = h(a, b)$
- ...

- The root hash is a hierarchical digest of entire set
- [Merkle]



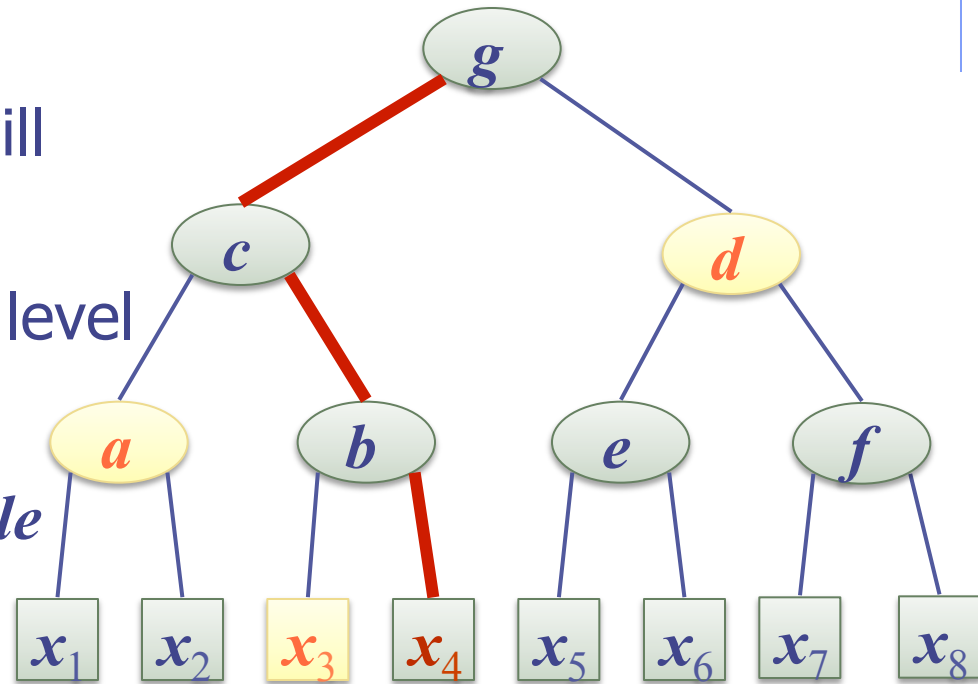
Hash Tree Verification

- Assumptions
 - Collision resistant hash function
 - Root hash is known
- Membership proof of an item
 - path from the item to the root (L/R sequence) plus hash values of sibling nodes
 - logarithmic size
 - logarithmic verification time
- Example
 - $g = h(h(a, h(x_3, x_4)), d)$



Proof intuition

- In order to provide a verifying proof for a different leaf element, the adversary will have to break collision resistance in at least one level of the tree
- *This happens with negligible probability*



Recap:

Solutions to cloud based storage

Approach	Client Space	Proof Size	Verification	Proof Computation	Updates?
Hash all	$O(n)$	$O(1)$	$O(1)$	$O(1)$	$O(1)$
MAC	$O(1)$	$O(1)$	$O(1)$	$O(1)$	NO
Merkle tree	$O(1)$	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(\log n)$

- Can you make everything constant?
- *Impossible under certain assumptions*